Fieldwork in Language Contact Situations

Claire Bowern

April 3, 2008

1   Introduction

Few communities are wholly linguistically homogeneous and completely isolated from
their neighbors. There is great variety in the extent of language and dialect contact in
different communities, and therefore consideration of language contact is important
in fieldwork and language documentation. However, fieldwork focused on individual
speech varieties (or languages) has tended to see multilingualism as a problem rather
than an opportunity, as a source of contamination of the data under consideration
rather than something to study in its own right (see, for example, Vaux and Cooper
1999:8, and the argument in Aikhenvald to appear). Such a view is understandable
in the context of the wish to describe a single standard linguistic variety and to
produce materials which are representative of the language as a whole. Trying to get
an accurate picture of language contact in a community may also make the research
project considerably more complex; conversely, excluding contact means excluding
potentially relevant data.

   While there are numerous ways in which aspects of language contact impinge on
fieldwork and language description, in this chapter I concentrate on three different
ways in which language contact and multilingualism are relevant to fieldwork (and

vice versa). The first is the question of what to study when a linguist goes to the field. In §2 I consider definitions of contact, typical contact situations, and ways of discovering if the the field site involves language contact. The second aspect of fieldwork and contact, in §3, concerns what is studied in a language contact situation. The third, covered in §4 and §5, involves the effects that linguistically diverse speech communities have on fieldwork. I take examples from both extreme contact, such as community-wide multilingualism, and less extreme examples. Finally, fieldwork in a language contact situation requires certain methods and techniques. These include (but are not limited to) data management techniques, stimulus materials, and coding of information.

Throughout this article, I stress the need for the linguist to have an understanding and knowledge of the social situations at work in the community under study, such as demographics and history. That is simply because any linguistic claim about language contact reduces to a claim about social behaviour of speakers.

2 Defining and diagnosing language contact in the field

I take a broad definition of the term "fieldwork". I assume that fieldwork is any type of linguistic data gathering where the linguist uses information from a pool of speakers interacting with each other in their usual environment. This definition includes immigrant communities (for example, the Vietnamese community in Houston) but does not include field methods classes as fieldwork, nor working with a single speaker in a university setting (see Bowern 2008, Hyman 2001). The definition of 'field' and 'fieldwork' becomes important when considering language contact, since it is only in a situation where the linguist may observe the different factors which make up language contact that progress can be made.

2.1 Defining contact and fieldwork

At the most basic level, all linguistic interaction is "language contact", albeit between extremely similar grammars. That is, speakers are exposed to many varieties of their language which differ in small ways from their own grammar. However, we use the term language contact to refer to situations where groups of people who speak very similar varieties are in contact with people who speak rather different varieties (cf. Thomason and Kaufman 1988, Thomason 2001:2). That is, there is more than one speech variety in use.[1] Fieldwork in such an area simply involves going to an area where contact occurs, beit in a large city or a small remote community. That is, I do not ascribe to the "Indiana Jones" fieldwork model (Bowern 2008:13–14) where a linguist has to go to a remote (and preferably dangerous) part of the world in order to have the work count as "fieldwork" (see also Hyman 2001) but equally studying

---

[1]For our purposes, a broader definition is more useful than a narrow one, since we are considering the situations in which one language may influence another, and how that relates to data gathering in the field. It is not necessary to quibble over whether the use of Latin in mediaeval Europe constitutes language contact, when the contact does not involve two groups of speakers with different first languages, one Latin, one vernacular.

a single speaker outside their regular patterns of language use will not allow much insight into those patterns. Therefore I am assuming a situation where the linguist has travelled to a community of speakers, not one where a single individual is being studied outside of their regular social networks.

Language contact is not, of course, a homogeneous phenomenon. Contact may occur between languages which are genetically related or unrelated, speakers may have similar or vastly different social structures, and patterns of multilingualism may also vary greatly. In some cases the entire community speaks more than one variety, while in other cases only a subset of the population is multilingual. Lingualism and lectalism[2] may vary by age, by ethnicity, by gender, by social class, by education level, or by one or more of a number of other factors. In some communities, there are few constraints on the situations in which more than one language can be used, while in others there is heavy diglossia, and each language is confined to a particular type of social interaction. An often-quoted example is the situation of various varieties of Arabic, or the status of French in West Africa or Swahili in Tanzania. In some parts of Indigenous Australia, in contrast, it is not uncommon for several languages to be in use at once, both through code switching and by different speakers speaking different languages (Evans 2001, Heath 1978).

There may also be different outcomes of language contact, ranging from limited lexical borrowing to pervasive metatypy (for the term, see Ross 1996, 1997). Thomason and Kaufman (1988:50, 74–75) diagram various possible outcomes of contact, categorised according to degree and intensity of contact, along with other factors. Thomason (2001:129–153) gives seven ways in which language contact may lead to language change. There are also discussions of language contact which play down the role of contact-induced change in language history, for example by arguing (e.g.

---

[2] That is, how many languages or 'lects' (varieties) of a language a person controls.

King 2000) that syntactic effects of language contact are caused by lexical borrowing (that is, the borrowing of syntax along with lexical items). Some discussion of these points of view can be found in Bowern (forthcoming) and Sankoff (2001). A fieldworker going to a language contact area should be familiar with this literature (and will probably be able to add to it!).

While there are a great number of different language contact situations, a few come up frequently in areas where linguists do fieldwork. One is dialect contact, for example between standard varieties of a language and regional varieties (e.g. in France or the Arab world). Such situations can exhibit strong patterns of diglossia. Linguists need to be aware of this because the situation in which they work tends to favor production of standard forms rather than nonstandard ones. For example, field workers working on Arabic frequently report difficulties in establishing work patterns where local (rather than standard) forms are elicited. Abbi (2001:Ch 7) mentions similar effects in parts of India.[3]

Trade or work languages are also often found, where the minority language is spoken within the community or at home and one or more sectors of the population conduct business in another language. In parts of southern Namibia, for example, Khoekhoegowab (also known as Nama/Damara) is the first language of many, but business, shopping, and so on is usually conducted in Afrikaans (and increasingly in English), and schools are English-only after grade four (see further Maho 1998).

A further type of language contact involves exogamous communities where more than one language might be used within the community because its members come

---

[3]For the purposes of studying contact in fieldwork, we need not make a difference between contact between languages and contact between standard and nonstandard varieties. The degree of difference between the various varieties makes the diagnosis of contact more or less easy, but it does not seem to be the case that degree of relatedness has an effect on degree of contact (Thomason and Kaufman 1988). Rather, the degree of contact and transfer depends on characteristics of social interaction and the type of social network, rather than the degree of relatedness between the varieties.

from different areas (see, for example, Aikhenvald (2006) for Amazonia, Stanford (2006) for Sui and Heath (1978) for Arnhem land, all areas where linguistic exogamy is practised). In northern Australia multilingual communities like Roper River, Milingimbi, Maningrida and Wadeye (Port Keats) arose from the settlement of various groups on missions, cattle stations and government depots. In some parts of the world, populations of long-term refugee camps have created their own lingua francas. The converse of such communities where exogamy leads to multilingualism is an endoterogenous community which maintains its own language for the purposes of excluding outsiders. Angloromani and Media Lengua (Muysken 1997, Winford 2003) are examples. In this case, extensive contact and transfer from majority languages may occur because the in-group language is learnt by adult second language learners.

Finally, fieldworkers particularly often work in endangered language communities where language shift is in progress. Such communities may exhibit "last speaker" effects of various types. For example, Thurgood (2003) reports the rapid restructuring of causative marking in the last generation of speakers of Anong. A similar phenomenon is often called "young people's varieties" and is quite common among the indigenous languages of Australia, where they are still being learnt by children (Langlois 2004, Lee 1987, Schmidt 1985). In §4.4 below I outline some of the ways in which these varieties may affect fieldwork. See also Romaine (this volume).

2.2 Symmetric and asymmetric multilingualism

Another potential point of variation concerns the patterns of multilingualism[4] within the community. Put simply, does everyone speak all the varieties that are used in the community? If not, who speaks what? What are the conditions under which each

---

[4]Note that throughout this article "multilingualism" implies competence in more than one language (that is, it is a cover term for non-monolingualism); I am not distinguishing bilingualism from multilingualism here.

language will be learned and used?

Truly symmetric multilingualism is sometimes argued to be quite rare. That is, the argument goes that in cases where the whole community speaks more than one language, the multilingualism is redundant and that some point it becomes unstable and the language shift occurs. These arguments are based on immigrant communities and heritage language speakers in places like the US, where we see a three generation pattern of shift from bilingual speakers to monolingual majority language speakers (see, for example, Fishman 1991). However, there are other situations where community wide multilingualism has, at least historically, proven to be stable over a longer period. One is the case of standard language diglossia versus regional varieties, such as Italy or Germany. Another is the case of small exogamous communities where there is a social expectation of multilingualism (see Aikhenvald (2004) for examples from Amazonia, and Heath (1978) for a discussion of an Australian case).

2.3 Endangered (Indigenous) language communities

One very common type of language contact situation is where a minority language, and often an endangered language, is spoken within a larger community. This language may or may not be the majority language within a specific speech community. For example, Khoekhoegowab is a minority language within Namibia, but it is the majority language in a number of villages in the south of the country. Other languages may be minority languages even within their own communities, such as Yan-nhaŋu at Milingimbi, where there are about 15 speakers and 100 community members out of a total community of about 800. Of course, not all Indigenous languages are endangered, and not all endangered languages are spoken by Indigenous people.

You will need to work out who speaks what language, and whether they speak it as a first language or a second language (or a later language). It is useful to find out

how exactly speakers learn such languages. It could be they picked them up as they grew up, with both languages be used in particular circumstances. Or, they might only have learnt another language with the school.

I have argued elsewhere (Bowern 2008) that Indigenous languages do not require special treatment in fieldwork; however endangered language communities and communities where language shift is in progress do require particular skills of the fieldworker. Some of these are relevant to language contact; others relate to the causes of language loss and shift more generally, such as poverty and community fracture. For more information on endangered languages, see Tsunoda (2004) and Grenoble and Whaley (1998).

2.4 Community/heritage languages

Another language contact situation concerns community languages, also known as heritage languages. These are languages of immigrant community groups, such as German or Finnish in the USA, Bengali or Balochi in the UK, and Moroccan Arabic in France (compare Clyne 1991, 2003). I mention heritage language groups as distinct from indigenous groups because they often treated differently in public policy (for example, in the English-only movements in the USA), and because there is often a standard language spoken outside the immigrant community which may serve as a point of comparison. Speakers may also look to that as a standard and may or may not consider their own variety to be different from it.

The dynamics of language rhetoric and shift are also distinct. For example, Indigenous groups are often told to 'modernise' their cultures by switching to the dominant language, whereas immigrant groups are usually told to 'assimilate'. Heritage language groups may have rather different views of their own variety vis-à-vis the standard. If they consider their own variety to be different, it may be that they

consider it a 'corrupted' variety of the language (and will want to teach you the standard). Conversely, if the population has been isolated from their country of origin for some time, they may consider their own variety to be more 'pure' or 'archaic'. Alternatively, there might be no perception that the variety is different from the standard.

2.5   Is your field site a language contact situation?

Unless you're doing monolingual elicitation (see e.g. Everett 2001), some language 'contact' is involved. Doing translation or fieldwork through another language will have an effect on your data. If you rely entirely on elicitation, you will bias the attested structures to ones that have easy translation equivalents in the language, and you'll probably only get the structures you think to ask about. Conversely, using only conversation will net you the most frequent structures but is unlikely to give you various other topics. For further discussion, see Bowern (2008:115ff).

However, I assume you are aware of the effects of the effects of a contact language in doing fieldwork, and we are talking about language contact which is not caused by the methodology of the linguist. Usually, it will be fairly obvious whether language contact is currently part of your field site, although it will probably not be clear what type of contact is occurring, who is participating, and to what extent.[5] From the very beginning of your fieldwork, you should be on the lookout for differences between speakers. There might be generational differences, gender differences, or other differences (although of course there may be other reasons for such differences other than language contact). You should also ask about language attitudes and about the language situation in the community. You may be able to get information about the demographics of the site and the wider area, which in turn may provide

---

[5]Furthermore, language contact may leave its mark on a language but the populations may no longer be in contact with one another.

information about likely conduits of language contact. Don't underestimate the utility of incidental observation as a source of topics for more detailed investigation by other methods.

Be aware however that just because an area involves because of more than one language it does not mean that there is necessarily a great deal of language contact. For example, Houston (a city of about 5 million people) has a very diverse population, with speakers of many different languages. We might therefore assume that it is a good place to study language contact. Further study would show, however, that while Houston is diverse overall, individual neighbourhoods tend to be very homogeneous, and so language contact is not nearly as great as one might suspect if one looked at the figures for the metropolitan area as a whole.

3 What to study

There are many possibilities for studying language contact and its results in different communities. We know that language contact can affect all levels of language (Thomason and Kaufman 1988), and any of these areas may be usefully studied.

The language contact may be an incidental part of the fieldwork. That is, you may be studying a language which also happens to be spoken in a language contact situation. Some of the issues that fieldworkers in such a situation should be aware of are listed in §3.1 below. Alternatively, you could study the linguistic results of language contact. That also brings up a particular set of issues, some of which are discussed in §5.

3.1 Working on a single language in a contact community

Traditionally, linguists have avoided working in multilingual areas where there is an option to work in a more homogeneous community. Handbooks often suggest trying to work with monolingual consultants were possible, to avoid the possibility of getting 'contaminated' data (Kibrik 1977, Vaux and Cooper 1999:8, Abbi 2001). However, if the aim of fieldwork is to produce an accurate documentation of a linguistic variety, as it is spoken by a group of speakers, language contact phenomena are a part of that and should be documented as well. To ignore code switching, borrowing, and other contact phenomena is to produce an artificial description of that language.

It is, of course, possible to work on a single variety. However you are likely to miss things (in all areas of the study), if you are unaware of the wider situation. If the contact has changed in the period between previous work on the language, you may find differences between your data and earlier sources, which might otherwise wish to ascribe to errors made by one on other of the linguists. You should compare your data with data outside the immediate community; this will allow you to gauge

the effect of language contact in the area. For example, a study of the differences between Dawkins (1916) and Greek spoken in Greece would show that the Asia Minor Greek varieties described by Dawkins vary in systematic ways from those spoken in mainland Greece. However, be wary of attributing all differences to language contact; it is tempting to see contact as the cause of all differences. In practice, cause can be very difficult to ascribe with certainty. For example, it is obvious that the syntax of Texas German[6] differs in systematic ways from Standard High German, and that many of the features which distinguish Texas German from Standard German are ones which Texas German shares with English. However, given that Standard German was only one of the inputs to Texas German (along with a number of regional dialects which differed markedly from Standard German), Standard German did not change into Texas German. For further discussion of the interplay of borrowing, drift, and change, see Jones and Esch (2002). For more information on Texas German, see Boas (2003), Salmons and Lucht (2006).

3.2   Working on the contact situation

Classic and more recent studies of language contact areas and their linguistic results include Dawkins (1916) for the Greek spoken in (Turkish-dominant) Asia Minor, Gumperz and Wilson (1971) for Kupwar village and the convergence of three languages from different families, Aikhenvald (2006) for the language contact situation in Amazonia, Nurse (2000) for Daiso (an area between Kenya and Tanzania), Heath (1978) for Arnhem Land, Ho and Platt (1993) for Singapore English, and many others. All of these studies were based on fieldwork, using various techniques. The data for Nurse (2000) began as a linguistic survey, while Aikhenvald (2006) is based on detailed work, initially with a single community in the region.

---

[6]See `tgdp.org`.

One way to approach such a study would be to take an anthropological linguistic perspective of the contact area and define the contact interaction itself as the focus of study. Working on the contact situation and its linguistic results can be framed in terms of a number of questions. What languages are used in the community? Who speaks what language? How well do they speak those languages? Who do they talk to, and what language do they use when they talk to them? That is, when are the different languages used? When people communicate, how do they utilise each of these languages? Do they codeswitch? What led to this situation? Why are there multiple languages in use in the community? How do community members acquire the different varieties in use in the community? Does everyone have the same degree of multilingualism, or is it asymmetric? What factors govern who is multilingual? Do all sections of the community participate in the multilingualism, or just some?

Working on a complex contact area requires preparation and it may take some time to begin to unravel the threads of the area. It might be necessary to have competence in several unrelated languages before seriously starting work. It might be necessary to work with a diverse cross section of the community in order to find the relevant variables to focus on. Such field sites should not be considered short-term, 'quick studies'. A superficial familiarity with the community is bound to result in bad generalisations.

Aikhenvald (to appear:1) presents the study of language contact in terms of the description of the multilingual competence of a single speaker. That is, she frames the discussion in terms of describing the social and linguistic competence of a speaker with multiple grammars at their disposal. It is, of course, possible to take the monolingual linguistic competence model and adapt it to a multilingual person. There is some work of this type, including Halmari (1997) (see also Roberts [this volume]). However, most studies of contact take either a community-oriented view or a historical one, and

discuss the situation not in terms of competing (or complementary) grammars within individual speakers but as a set of community resources which influence each other.

3.3 Working on a contact variety

A further approach to language contact is the study of a single contact variety, such as a pidgin, creole or mixed language (for the terms see Holms [this volume]). I would argue that from the fieldwork perspective, work on such languages is no different from that described in §3.1, that is, working on a single language in a contact situation. Contact varieties can (and should, I would argue) be approached with exactly the same tools that are used for any other type of linguistic fieldwork.

Resist the temptation to describe the language in terms of the other languages that went into its genesis. That is, it is very tempting to describe a contact variety such as Mednij Island Aleut in terms of Russian and Aleut (the two languages which contributed to the variety), or to describe Young People's Dyirbal in terms of the differences between it and Traditional Dyirbal (Schmidt 1985). However, in many cases, there will be features of the new variety which are not present in either of the contributing varieties. After all, Modern Bardi has considerably more features of polysynthesis than the Bardi of the 1920s, but Bardi has not acquired those features from English even though the primary language contact in that community has been with English. Moreover, just because a contact variety arose, it does not imply necessarily that the contact situation is the same today. Kriol is a contact variety which arose in northern Australia about a hundred years ago, and most of the languages spoken by the community at Roper River who gave genesis to the creole are no longer spoken. These days, the main contact is between Kriol and English, not between Kriol and Indigenous languages (Harris 2004, 2007).

4 Linguistic and paralinguistic effects of language contact relating to fieldwork

4.1 Identifying cause and effect

Identifying cause and effect within data is known to be difficult, not only in linguistics. This is also true in establishing that particular shared similarities are due to language contact. After all, any given shared feature may not be due to language contact between those two varieties, but rather to retention of shared features (if the languages are related), borrowing from a third language, calquing, or chance. Frequently, in language contact areas, any feature which may be attributed to contact is done so. This in turn implies a particular view of language change and of language use in a community. Establishing the causes of change and the sources of particular constructions in contact areas is extremely difficult, and plausible analyses may be found which could cover several different situations. Therefore it is very important to think about the evidence for a claim of contact-induced language change.

4.2 Variation in data

Multilingual communities can exhibit a great degree of variation, both within the data from a single speaker and throughout the community as a whole. This can trip up the unsuspecting fieldworker who expects to be describing a single cohesive variety. (Of course, there is variation in all language, but the variation in complex contact areas with multilingual speakers can be especially clear, since speakers can draw on multiple speech varieties for different purposes and the fieldworker does not have access to those nuances.)

A linguistic variable is any linguistic item which has different realisations. The different realisations may be conditioned by any number of factors, including the

age of the speaker (speakers over 50 may have a preference for one pronunciation of an item, whereas school-age children may systematically use another realisation of that item), the social class, race or ethnicity, gender, level of education, or degree of familiarity with other languages. That is, contact-induced variation cross-cuts the other types of variation which are found in communities, no matter whether language contact is present. Variation may be stable (such as [æks] ∼ [ask] for English 'ask') or it might be indicative of a linguistic change in progress.

4.3 The language under study

In areas where extensive codeswitching is normal, it might be extremely unnatural for someone to talk in a single language. This makes it very difficult for them to produce extended chunks of speech without codeswitching. It leads to much greater self-monitoring and may interfere greatly with grammaticality judgments. That is, the speaker is so concentrated on the production of speech in a single language that they produce sentences which are stilted and unnatural. The best way to deal with a situation like this is not to worry too much about the codeswitching, but to check the language later with another speaker. Speakers will often correct codeswitches in such contexts. This gives you valuable data not only about what forms speakers feel belong to one language or another, but also what the equivalent non-switched forms are.

Similar issues may occur if the language is not used regularly, for example if the language is highly endangered and most of the population has already shifted to another language. If the person is not used to producing sentences in that language, you may get answers which include words and syntax from several languages. Some discussion of working with semi-speakers can be found in Bowern (2008:137–139), including how to encourage people who may not have spoken the language for a long

time.

Speakers may be unable to tell in some situations what forms belong to which language. Although we learn from introductory linguistics onward that speakers intuitively know what is well formed in their language and what is not, in practice, speakers cannot always assign the correct form to the correct language. Judgements about which word belongs to a particular language may vary considerably. For example, if you ask a speaker of English if [ˈdʒəntə] 'junta' is an English word, there are several possible responses. On the one hand, it is a loan from Spanish, and therefore not an English word in the sense that 'stone' is an English word. On the other hand, *junta* is not pronounced [ˈdʒəntə] in Spanish; therefore it's not the same as the Spanish word with the same orthography. I have witnessed this problem frequently in my own fieldwork in eastern Arnhem land, especially in the context of extensive elicitation and direct questioning. Occasionally in fieldwork involving lexical elicitation I have had speakers produce the target word in English, with the phonology of the target language. That is, I have asked for the word *sort* and received the answer [sɔrtʰ].

In extreme cases, you may get data in the "wrong" language. For example, McDonald and Wurm (1979) is a grammar of the Wangkumara language and was documented from a single speaker. However, the speaker called the language Garlali, which is the name of the language spoken to the east of Wangkumara, and in fact, the authors were under the impression (as far as proof stage of the grammar) that the language theyw ere documenting was Garlali.

4.4   Which language to record, which variety to document

Language communities with extensive variation present questions for documentation. In some types of language contact community, it is common for young people to have

17

a very different way of speaking from their parents and grandparents. These "young people's varieties", as mentioned above, are quite common in Australia (cf. Lee 1987, O'Shannessy 2005, Schmidt 1985 among others). Of course, both the older more traditional way of speaking and the young people's speech are extremely interesting and could be the target of a documentation project. However, the variety you choose to do fieldwork on has consequences for the project. The community may feel that the young people do not speak properly, and that if you use their variety as the basis for the documentation you will not be recording the right version of the language and you risk alienating some of the most powerful people in the community. On the other hand, if you pick the more prestigious variety and then base language learning or other documentation materials on it, you are likely to alienate the target audience for the materials. In such cases, let the community may guide you, although of course this supposes that the 'community' is in agreement, which is unlikely to be true.

Some thought should be put into documentation materials in language contact areas. Linguists tend to talk about 'giving back' to the community (e.g. Bowern 2008:Ch. 14, Crowley 2007:34) and frequently this takes the form of learner's guides, school project materials, oral history books, language software or other materials designed to promote or showcase the language. However well-meaning such proposals may be, in some areas they might be greeted as inflammatory or patronising. Consider the response that writing street signs in African American English would have in Harlem, for example, especially if it were promoted by university academics.

4.5 Language attitudes

Even if the primary object of study is a single variety in a contact situation and not the language contact itself, language attitudes will likely have a great effect on both the data and the circumstances of the fieldwork.

First, if you are working with people who are fluent in both the local regional language and the standard language, you may only get data from the standard as the default response. This is especially true if you are a native speaker of either the standard variety or speak a different variety of the language you are doing fieldwork on. Because it is usual in such circumstances for people to speak the standard language, it may be quite difficult to elicit anything else. Furthermore, certain types of linguistic techniques tend to promote particular linguistic interactions (and it is necessary to be aware of them even if they are not what you are studying). Elicitation is a type of formal or semi-formal interview, and formal contexts tend to elicit particular varieties in areas of high diglossia. Vaux and Cooper (1999) have some suggestions for encouraging discussion of low-prestige varieties. Recruiting local research participants to do interviews will also minimise this problem.[7]

The race, ethnicity and/or gender of the participants in the conversation may affect the data. It might be hard to overcome tendencies to talk in one language to someone from that group. There is a reported issue in language documentation that elders who are used to not talking their language in front of younger community members find it difficult to break those habits. I once had a conversation with someone where I spoke his language and he spoke English. At one point he apologised that he had never spoken his language in front of a white person before and found it impossible to do so.

In a highly multilingual community, it may be that the medium you work through as a field language will harm your ability to get data in your target language. This is especially the case where there is the default language which overrules the use of other languages or varieties. For example, if there is a standard language and you are trying

---

[7]Recording speakers without their knowledge is sometimes also suggested as a way to minimise the effects of a recording device on fluency, but since that is unethical

to study a local variety. There are numerous cases of field workers reporting that they had to deny knowledge of major languages (which they in fact spoke fluently) in order to have any access to the smaller languages.

Aspects of language politics and language purity may also come into play. In §4.3 I mentioned the case of a man who gave linguists a different language from the one they thought they were working on because of a personal preference for speaking one language over another. In other cases, speakers may reject all words which are shared with other languages they know as being 'not proper X' or 'borrowed from Y', even if those words are in general use or are not, in fact, borrowed. In one of my field sessions, one person rejected all present tense forms from a particular verb conjugation, on the grounds that they were "borrowed" from the local *lingua franca*. In fact, that form is common to a number of related areas in the region and the shared forms are retentions from an earlier Proto-language, and not borrowed forms. Moreover, the morphemes in question are used in different ways in the two varieties.

5 Field techniques

Working in multilingual areas makes it especially important to have good data gathering procedures. It is also especially important to document field notes and recordings. If more than one language is being worked on at a time, an undocumented collection can make it extremely difficult to sort out what language is what, especially if there is more than one undocumented language in the sample.

5.1 General field techniques

As I have stressed elsewhere in this article, many of the techniques used in field linguistics more generally are also applicable when working on language contact (see Bowern (2008), Crowley (2007), Gippert *et al.* (2006), and Newman and Ratliff (2001) for some advice). The linguist should take nothing for granted, but test all assumptions. It is often necessary to keep several conflicting hypotheses in mind simultaneously. Work with multiple speakers, compare results, be aware of variation and homogeneity in the data. Keep good records. Learn to speak the language under study and use the intuitions you gain from familiarity with your data for more data testing. Look for patterns rather than individual features. Field sites with extensive linguistic variation are extremely confusing when the fieldwork is at an early stage. It is difficult enough to do fieldwork on a single language; that complexity is compounded when there is need to work not only on several languages at once, but also on the relationships between those languages. Many sites are not so complex, thankfully.

5.2 Ethnographic methods

Most of your information on language contact in a field site will come from three sources. The first is what you can infer from the languages themselves, from data and techniques such as §5.3. The second is interviews and self-reports. Observation

is the third source of potential information.

It is fine to ask people about their views of language and when they use particular items. You can ask about impressions about who speaks the 'same' and 'differently', who uses particular varieties, which varieties are prestigious and which are stigmatised, and other information about language attitudes and perceptions of variation. Niedzielski and Preston (2003) is a discussion of folk linguistics and folk linguistic categorisations; Eckert (2000) is a very detailed example of the application of ethnographic methods in sociolinguistics (although not in a language contact area).

What speakers identify as a contact might not be true contact. That is, while speakers may be sensitive to the similarities and differences between the languages they speak, they probably won't know the history of language contact and there is no a priori reason to take their statements about what are and aren't borrowings at face value.

Observation is also a very powerful tool for investigating language contact. It is a core tool in sociolinguistics and should be used in documentary/descriptive linguistics, and even in fieldwork aimed primarily at theoretical topics. After all, the best way to find out what is interesting in a speech community (and worth testing further) is to do some exploratory work.

There is extensive discussion of observation and ethnographic methods in the sociolinguistic literature. Johnstone (2000), Meyerhoff (2006), Milroy (1980, 1987), Wardhaugh (1986) are good references to begin with for further information.

5.3   Grammaticality judgments, elicitation and stimulus materials

While ethnographic methods such as participant observation are invaluable for fieldworkers, direct questioning about linguistic data is also necessary. Sociologically oriented studies of language tend to play down the value of translation and grammat-

icality judgements (and the like) on the grounds that they are easy to manipulate unintentionally, difficult to quantify and subject to unreliability. However, like all methods, elicitation is dependent on the skills of the person using the method, and these methods work better in some cases than in others. Furthermore, direct translation or elicitation is not the only method for obtaining linguistic data. The methods considered in this section fall into three types: translation/elicitation, controlled tasks, and content checking.

In translation (or elicitation narrowly defined), the linguist asks the speaker of the language to translate sentences from the contact language into the target language. The translations are designed to give the linguist information about the different vocabulary and structures in the language. Problems with elicitation as the primary method for data gathering include biases of frequency (that is, elicited gives the linguist information about a) structures which are easy to translate, and b) structures which the linguist thinks to ask about), and the possibility of data contamination from unnatural translations (that is, the translation is a literal one, but which is not a felicitous expression in the language). Some of these problems can be avoided by being explicit about the task instructions, and by checking the answers with another speaker.

Controlled creative tasks and stimulus materials avoid some of the problems mentioned in connection with direct translation, because they do not bias word or construction choice. Examples of such tasks include retelling the *pear stories* (Chafe 1980), *frog stories* (Berman and Slobin 1992), segmenting colour wheels (Berlin and Kaye 1969), providing vernacular definitions or descriptions of items, and describing video clips, pictures, or objects. The results are then transcribed and translated into the target language and glossed for analysis. Such tasks are very useful in documentation because they allow the linguist to guide the topic and structures for description

without prespecifying them. However, it is difficult to get exactly parallel data this way.

The third method is content checking. That is, the linguist asks specific questions about constructed sentences or previously gathered data in order to find out more about the language. These questions might take the form of requests for grammaticality judgements ("is 'X' a good sentence?"), requests for more information about the meaning of a word or phrase, back-translation, or checking transcriptions and translations from earlier sessions. All of these methods are likely to produce variable answers and can serve as the input to further investigation.

Finally, consider the secondary uses of materials originally collected for another purpose. A well-annotated collection of data can be used for multiple projects. Once a set of sound files are segmented and tagged for a study on vowel length, for example, the same data can be used for consonant duration, or a formant analysis. A set of texts which have been edited to remove codeswitching for a published version also serves as the basis for a study of that codeswitching.

5.4 Metadata and annotation

You need some way of keeping track of all the data that you have recorded. Good data is made much more useful by good metadata (that is, data about the data itself; see Bowern (2008:56ff)). It is always true that annotated data is more easy to use than unannotated data. Consider the situation of a set of notes in an unknown language without translations versus one with the translations. Data from language contact situations is complex because it contains pieces from interlocking systems with different properties. Having good information about the data makes this easier.

Keeping track of linguistically diverse data can be challenging. Data points should be associated with the speaker (or by speaker categories, when trying to work out

what the patterns of variation are). That is, you should record where all your data comes from. This means you will also need good speaker metadata: at least age (or approximate age), gender, language background, class and occupation (if relevant), and clan (if relevant). You will need to know where and when each session was recorded, and who else was present.

Next, you will need to be able to find your data again, so you will need some way of keeping track of the topics you have recorded. If you use the same prompt materials with several speakers, it will be most useful if the data are transcribed in a format that allows direct comparison between the versions. Time-aligning transcriptions (that is, recording using a digital recorder and transcribing the recordings using software which allows the sound clip to be lined up with the transcription. There are several such free programs available.

6   Conclusions

In many ways, multilingual fieldwork is no different from fieldwork in a linguistically homogeneous community. Fieldwork in language contact situations is diverse, just like fieldwork everywhere. The best approach to a linguistically diverse fieldwork community is a diverse analysis toolkit which includes a variety of techniques coupled with good quality recording and annotation. Where feasible, such a toolkit includes observation and other ethnographic techniques, backed up with experimental and elicited data from a sample of the population.

References

Abbi, Anvita, 2001. *A manual of linguistic field work and structures of Indian languages*. München: Lincom Europa.

Aikhenvald, A Y, 2004. *Language contact in Amazonia*. Oxford.

—, 2006. Serial verb constructions in typological perspective. In *Serial Verb Constructions: A Cross-Linguistic Typology, 2006*, Cambridge University Press, chap. 1. pp. 1–68.

Aikhenvald, Alexandra, to appear. Multilingual fieldwork, and emergent grammars. In *Proceedings of the 33rd Annual Meeting of Berkeley Linguistics Society*, Berkeley, Calif: Berkeley Linguistics Society.

Berlin, Brent and Paul Kaye, 1969. *Basic color terms: their universality and evolution*. Berkeley: University of California Press.

Berman, Ruth and Dan I Slobin (Eds.), 1992. *Relating evants in narrative: a crosslinguistic developmental study*. Lawrence Erlbaum.

Boas, Hans, 2003. Tracing Dialect Death: The Texas German Dialect Project. In J Larson and Mary Paster (Eds.), *Proceedings of the 29th Annual Meeting of the Berkeley Linguistics Society*, Berkeley, Calif: Berkeley Linguistics Society, vol. 29. pp. 387–398.

Bowern, Claire, 2008. *Linguistic Fieldwork: a practical guide*. Basingstoke: Palgrave Macmillan.

—, forthcoming. Syntactic change and syntactic borrowing in generative grammar. In Gisella Ferarresi and Maria Goldbach (Eds.), *Handbook of historical syntax*, Amsterdam: Benjamins.

Chafe, Wallace (Ed.), 1980. *The pear stories: cognitive, cultural and linguistic aspects of narrative production*. New Jersey: Norwood.

Clyne, Michael, 1991. *Community languages: The Australian experience*. Hong Kong: Cambridge University Press.

—, 2003. *Dynamics of Language Contact: English and Immigrant Languages*. Cambridge University Press.

Crowley, Terry, 2007. *Fieldwork linguistics: a beginner's guide*. Oxford: Oxford University Press.

Dawkins, 1916. *Modern Greek in Asia Minor*. Oxford.

Eckert, Penelope, 2000. *Linguistic variation as social practice: the linguistic construction of identity in Belten High*. Blackwell Publishers.

Evans, Nicholas, 2001. The last speaker is dead; long live the last speaker. In Paul Newman and Martha Ratliff (Eds.), *Linguistic fieldwork*, Cambridge: Cambridge University Press, chap. 12. pp. 250–81.

Everett, Daniel, 2001. Monolingual field research. In Paul Newman and Martha Ratlif (Eds.), *Linguistic fieldwork*, Cambridge: Cambridge University Press. pp. 166–188.

Fishman, Joshua, 1991. *Reversing Language Shift: Theoretical and Empirical Foundations of Assistance to Threatened Languages*, vol. 76. Multilingual Matters.

Gippert, J., N. Himmelmann and U. Mosel, 2006. *Essentials of language documentation*, vol. 178 of *Trends in Linguistics, Studies and Monographs*. Mouton de Gruyter.

Grenoble, L.A. and L.J. Whaley, 1998. *Endangered Languages: Language Loss and Community Response*. Cambridge University Press.

Gumperz, J.J. and R. Wilson, 1971. Convergence and creolization: A case from the Indo-Aryan/Dravidian border in India. *Pidginization and Creolization of Languages*:151–167.

Halmari, Helena, 1997. *Government and Codeswitching: Explaining American Finnish.*. Philadelphia: John Benjamins.

Harris, J., 2004. Kriol – the creation of a new language. In Suzanne Romaine (Ed.), *Language in Australia*, Cambridge: Cambridge University Press. p. XX.

Harris, John, 2007. Linguistic responses to contact: Pidgins and creoles. In Gerhard Leitner

and Ian Malcolm (Eds.), *The Habitat of Australia's Indigenous languages: past, present, and future*, Berlin: Mouton. pp. 131–152.

Heath, Jeffrey, 1978. *Linguistic diffusion in Arnhem Land*. Canberra: Australian Institute of Aboriginal Studies.

Ho, Mian-Lian and John Platt, 1993. *Dynamics of a contact continuum: Singapore English*. Oxford: Oxford University Press.

Hyman, Larry, 2001. Fieldwork as a state of mind. In Paul Newman and Martha Ratliff (Eds.), *Linguistic Fieldwork*, Cambridge University Press, chap. 1. pp. 15–33.

Johnstone, Barbara, 2000. *Qualitative methods in sociolinguistics*. Oxford University Press.

Jones, M and E Esch (Eds.), 2002. *Language change: the interplay of internal, external and extra-linguistic factors*. Mouton.

Kibrik, Alexander, 1977. *The Methodology of Field Investigations in Linguistics: Setting Up the Problem*. Ianua Linguarum, The Hague: Mouton.

King, Ruth, 2000. *The lexical basis of grammatical borrowing*. Amsterdam: John Benjamins.

Langlois, A, 2004. *Alive and Kicking: Areyonga Teenage Pitjantjatjara*. Canberra: Pacific Linguistics.

Lee, Jennifer, 1987. *Tiwi today: study of language change in a contact situation*. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Maho, J. F., 1998. *Few people, many tongues: the languages of Namibia*. Windhoek: Gamsberg Macmillan.

McDonald, M A and S A Wurm, 1979. *Basic materials in Wangkumara (Gaḷali)*, vol. B-65. Canberra: Pacific Linguistics.

Meyerhoff, Miriam, 2006. *Introducing Sociolinguistics*. London and New York: Routledge.

Milroy, Lesley, 1980. *Language and social networks*. Oxford: Blackwell.

—, 1987. *Language and social networks*. Oxford: Blackwell, 2nd edition edn.

Muysken, Peter, 1997. Media lengua. In *Language contact: a wider perspective*, Amsterdam: John Benjamins. pp. x–y.

Newman, Paul and Martha Ratliff (Eds.), 2001. *Linguistic fieldwork*. Cambridge: Cambridge University Press.

Niedzielski, Nancy and Dennis Preston, 2003. *Folk Linguistics*. Berlin: Walter de Gruyter.

Nurse, Derek, 2000. *Inheristance, Contact, and Change in two East African languages*. Köln: Rüdiger Köppe Verlag.

O'Shannessy, Carmel, 2005. Light Warlpiri: A New Language. *Australian Journal of Linguistics*, **25**(1):31–57.

Ross, Malcolm D, 1996. Contact-induced change and the comparative method: cases from Papua New Guinea. In M. Durie and M Ross (Eds.), *The Comparative Method Reviewed*, New York: Oxford University Press, chap. 8. pp. 180–217.

—, 1997. Social networks and kinds of speech-community event. In Roger Blench and Matthew Spriggs (Eds.), *Archaeology and Language I*, London: Routledge, chap. 13. pp. 209–261.

Salmons, Joseph and Felecia Lucht, 2006. Standard German in Texas. In Linda Thornburg and Janet Fuller (Eds.), *Studies in Contact Linguistics: essays in honor of Glenn G. Gilbert*, New York: Peter Lang. pp. 167–188.

Sankoff, Gillian, 2001. Linguistic outcomes of language contact. In *Handbook of Sociolinguistics*.

Schmidt, Annette, 1985. *Young people's Dyirbal. An example of language death from Australia*. Cambridge Studies in Linguistics, London: Cambridge University Press.

Stanford, Jim, 2006. When your mother tongue is not your mother's tongue: Linguistic reflexes of Sui exogamy. *University of Pennsylvania Working Papers in Linguistics: Selected Papers from NWAV-34*, **12**(2):217–229.

Thomason, Sarah Grey and Terrence Kaufman, 1988. *Language contact, creolization and genetic linguistics*. Berkeley and Los Angeles: University of California Press.

Thomason, S.G., 2001. *Language Contact*. Edinburgh University Press.

Thurgood, G., 2003. The causatives in sun hongkai's anong: language death and rapid restructuring. *Proceedings of the 28th Annual Meeting of the Berkeley Linguistics Society*.

Tsunoda, T., 2004. *Language Endangerment and Language Revitalization*. Trends in Linguistics. Studies and Monographs, The Hague: Mouton de Gruyter.

Vaux, Bert and Justin Cooper, 1999. *Introduction to Linguistic Field Methods*. Munich: Lincom Europa.

Wardhaugh, R., 1986. *An Introduction to Sociolinguistics*. Blackwell Publishing.

Winford, Donald, 2003. *An introduction to contact linguistics*. Oxford: Blackwells.

CLAIRE BOWERN

Claire Bowern is an assistant professor of linguistics at Yale University. Her major research areas are historical linguistics and language documentation and description, especially relating to the Indigenous languages of Australia.

Contact details (Note: I am moving universities as of July 1).

Before July 1	After July 1
Department of Linguistics	Department of Linguistics
Rice University	Yale University
6100 Main St	370 Temple St
Houston, TX	New Haven, CT
77098 USA	06511 USA

Email: `bowern@post.harvard.edu`